

ABSTRACT

Method and framework for identifying optimal allocations of computing resources in a complex, distributed data processing environment. A plurality of server models are established, with each server model including one or more server nodes and each server node having an associated set of capacity attributes. Similarly, a plurality of service models are established, each service model including one or more service nodes and each service node having an associated set of demand attributes. The server models are defined with a layered relationship as are the service models. A node that is part of a model in one layer corresponds to a model in the next-lower layer. The invention generates optimized mappings of service nodes that are described in user-selected service models to server nodes that are described in user-selected server models, as a function of the associated sets of demand and capacity attributes.